

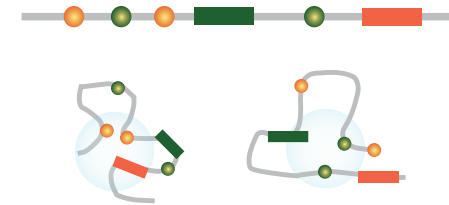
Chromatin structure and 3C-like data

Davide Baù & François Serra

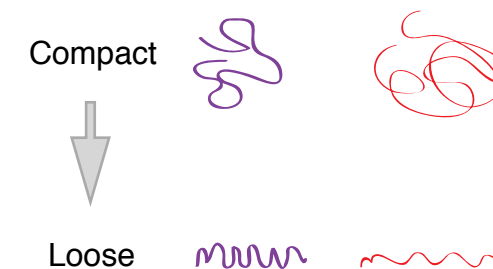
Genome Biology Group (CNAG)
Structural Genomics Group (CRG)

The role of chromatin structure

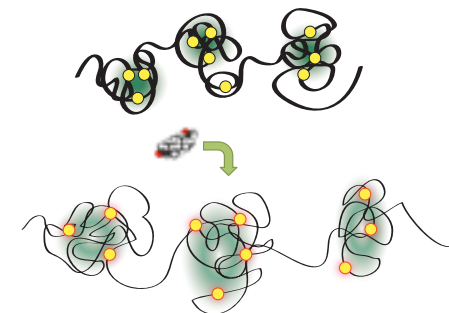
It can give insights into how distant genomic elements interacts with each other



It helps to understand the compartmentalization of chromosomes within the nucleus



It is essential to understand the mechanisms that regulate the cell



Chromatin definition

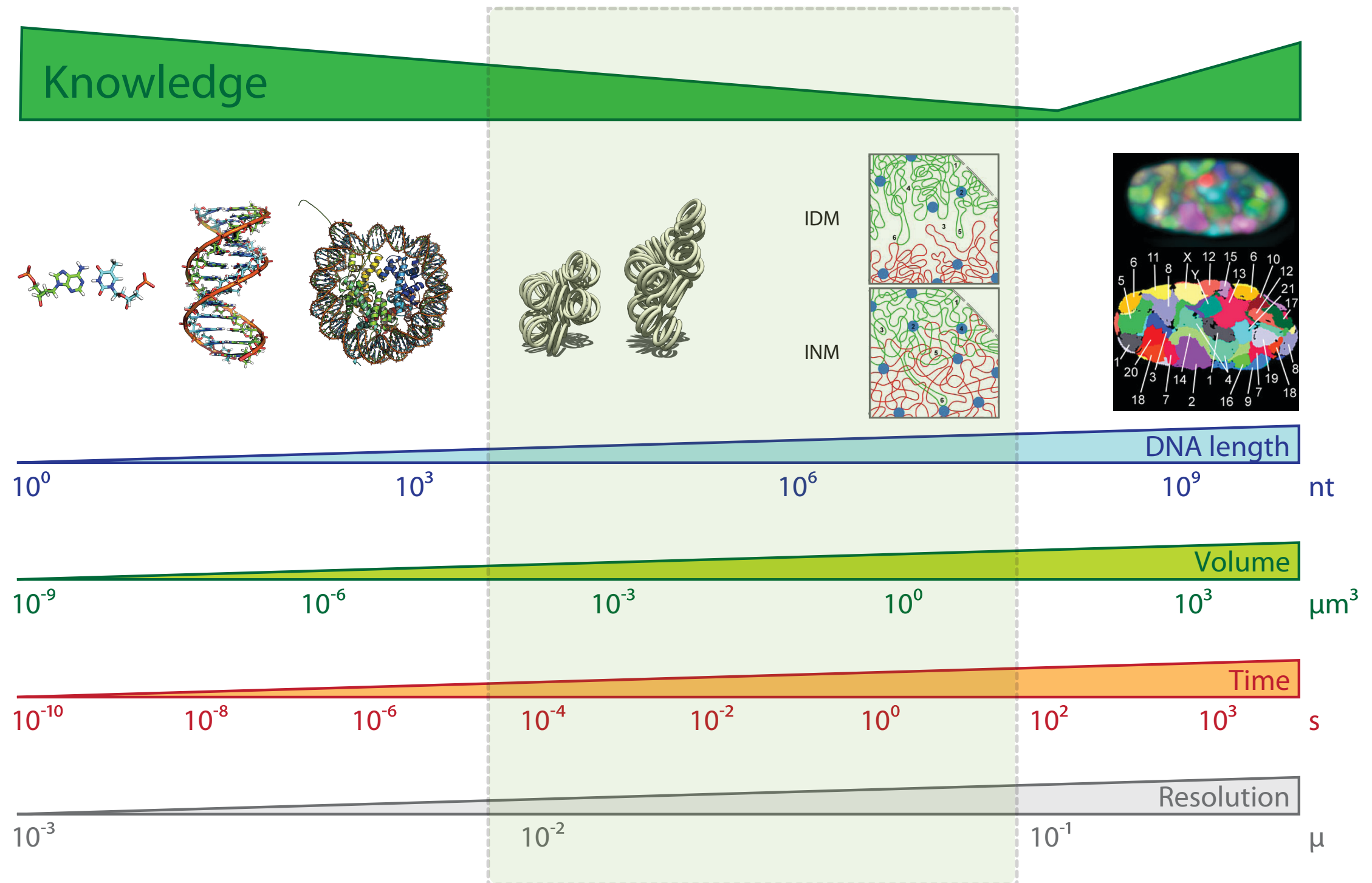
Chromatin is composed of DNA complexed with histones and other proteins

Chromatin formation enables the genome to be hierarchically packaged or condensed so that it can fit inside the nuclear space

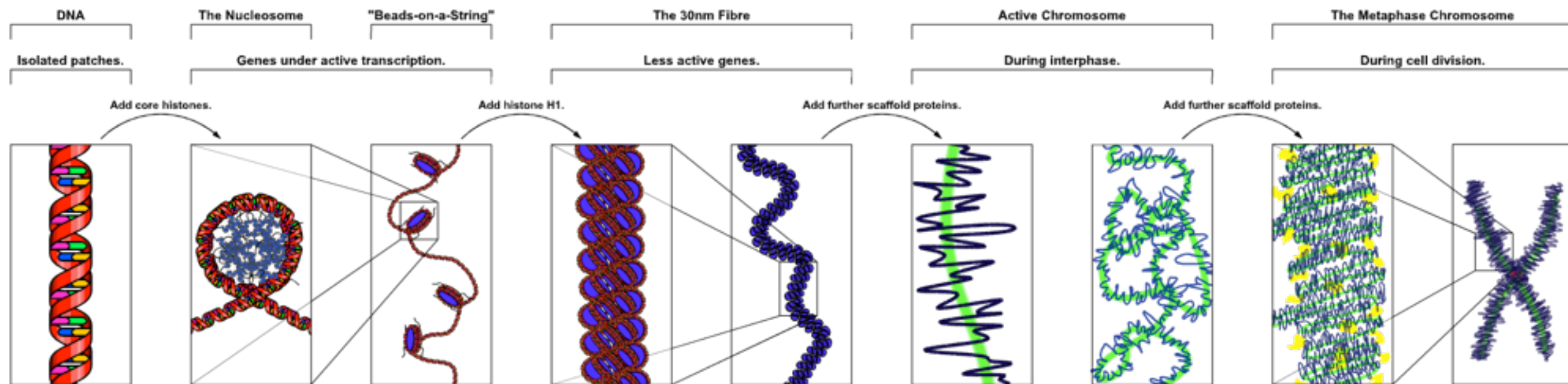
The compaction allows to modulate gene transcription, DNA repair, recombination, and replication

Chromatin structure is considered highly dynamic

Resolution gap



Chromatin structures

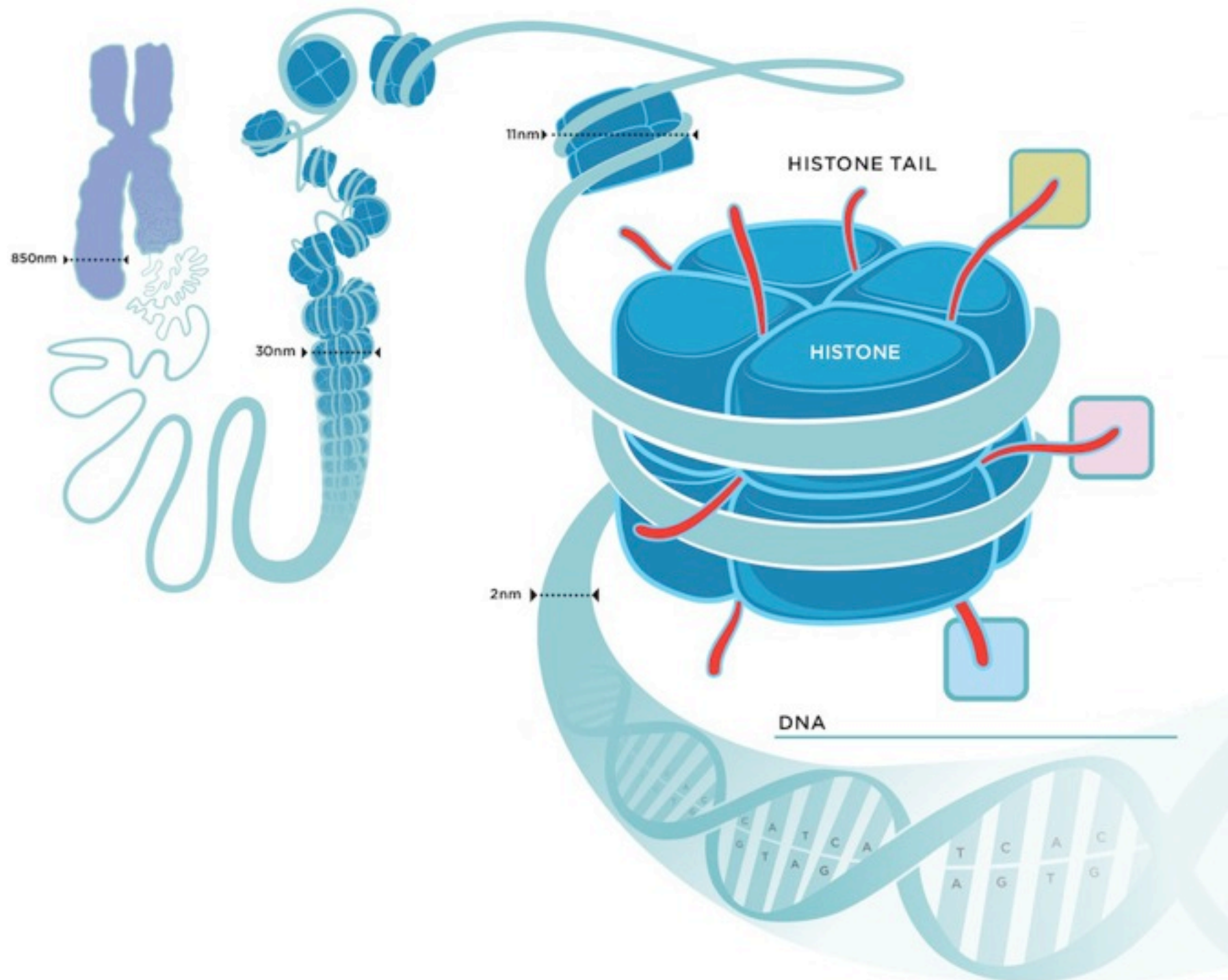


The nuclear organization of DNA

Chromosome

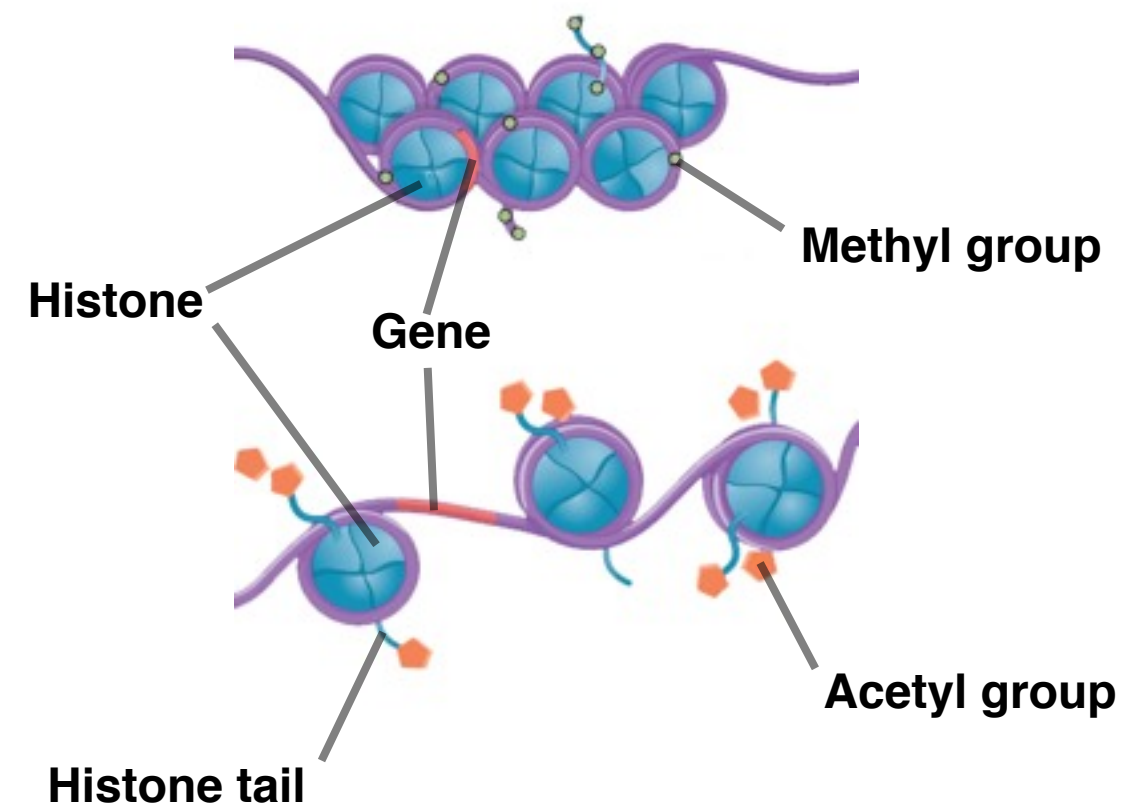
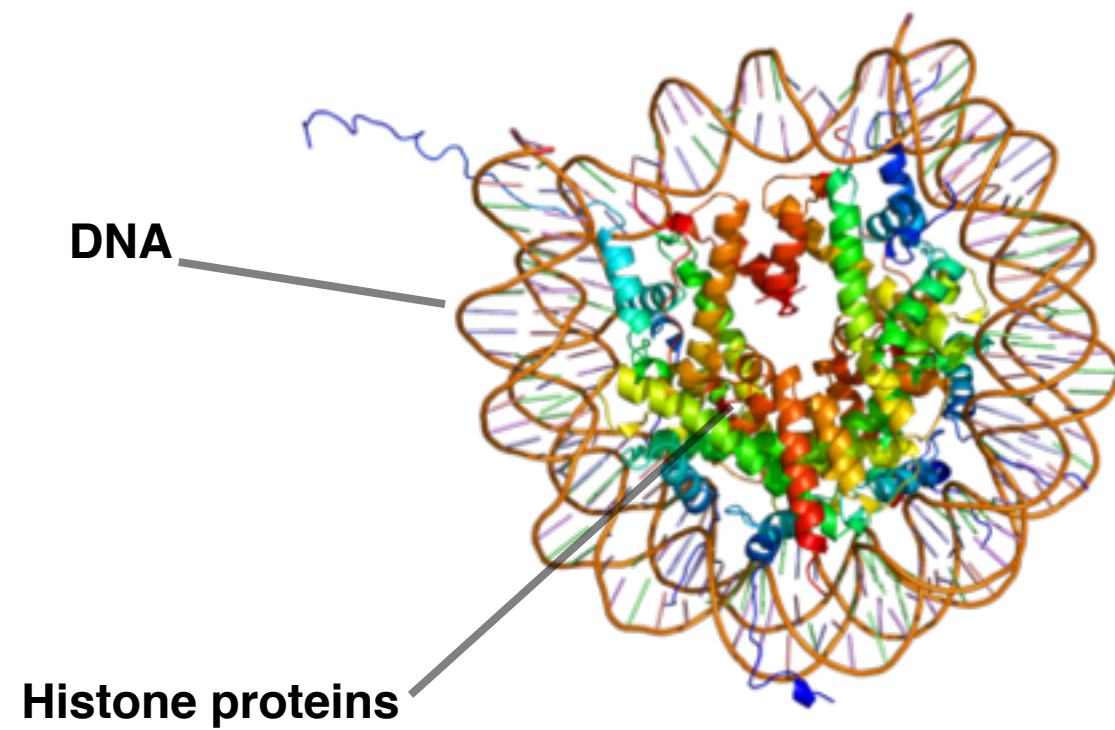
Chromatin fibre

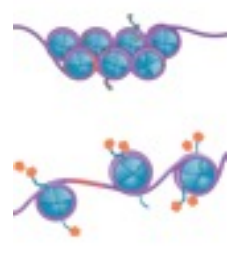
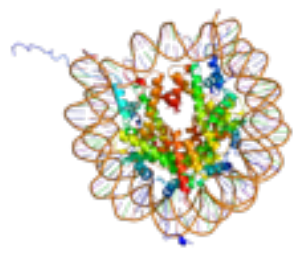
Nucleosome



Adapted from Richard E. Ballermann, 2012

The nucleosome



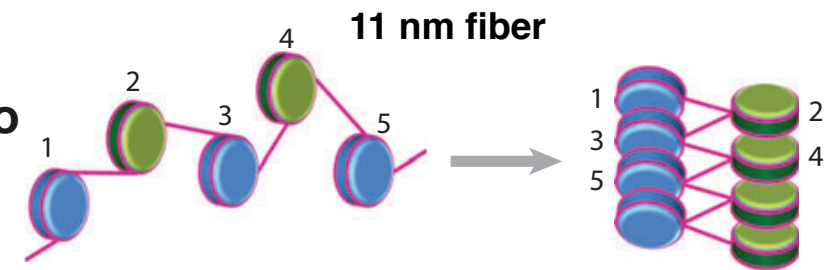


Histone modification effects

Type of modification	H3K4	H3K9	H3K14	H3K27	H3K79	H4K20	H2BK5
mono-methylation	activation	activation		activation	activation	activation	activation
di-methylation	activation	repression		repression	activation		
tri-methylation	activation	repression		repression	activation, repression		repression
acetylation		activation	activation				

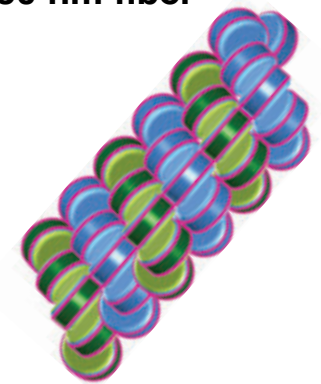
The chromatin compaction levels

Several nucleosomes in a row form what is often referred to as a beads-on-a-string fiber (the 11 nm fiber)



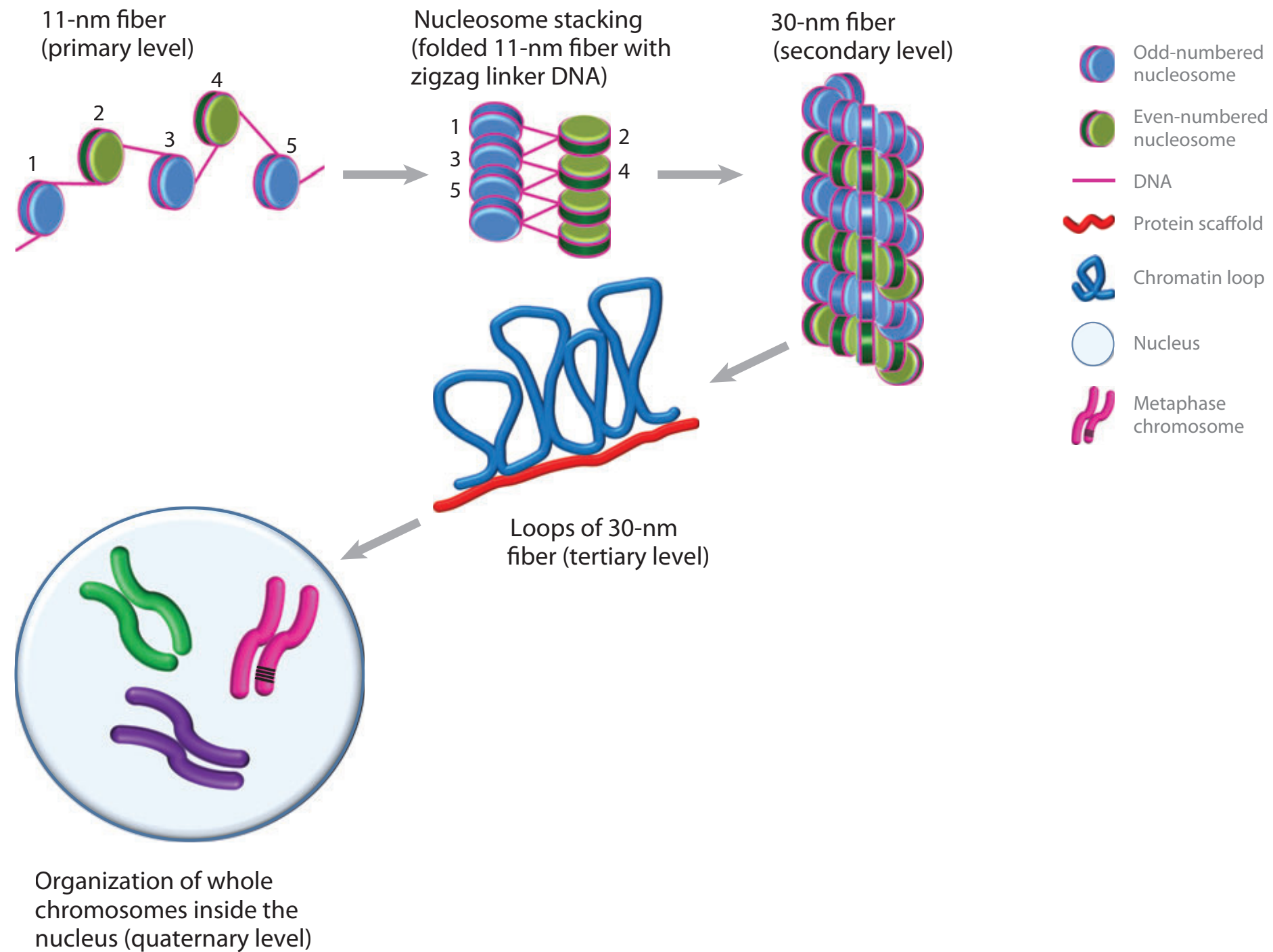
When histones H1 or H5, referred to as linker histones, are added to the 11-nm fiber, the condensed 30 nm fiber is formed

30 nm fiber



The 30 nm fibers form the next level of compaction by forming loops

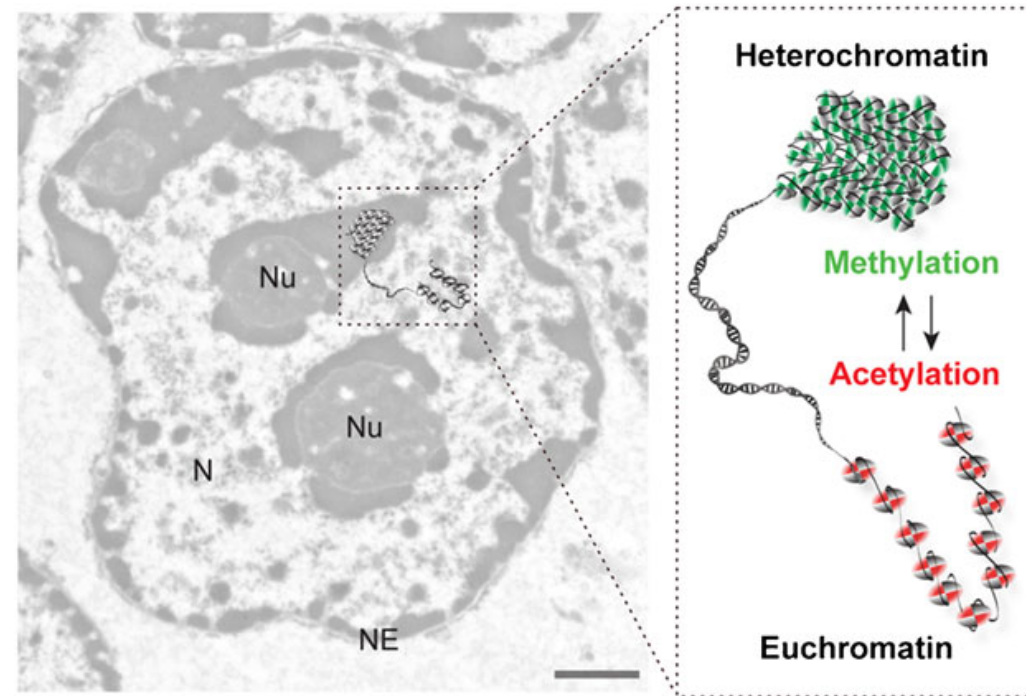
The chromatin compaction levels



Adapted from Annu. Rev. Genomics Hum. Genet. 2012, 13:59-82

Euchromatin and heterochromatin

Electron microscopy



Euchromatin:

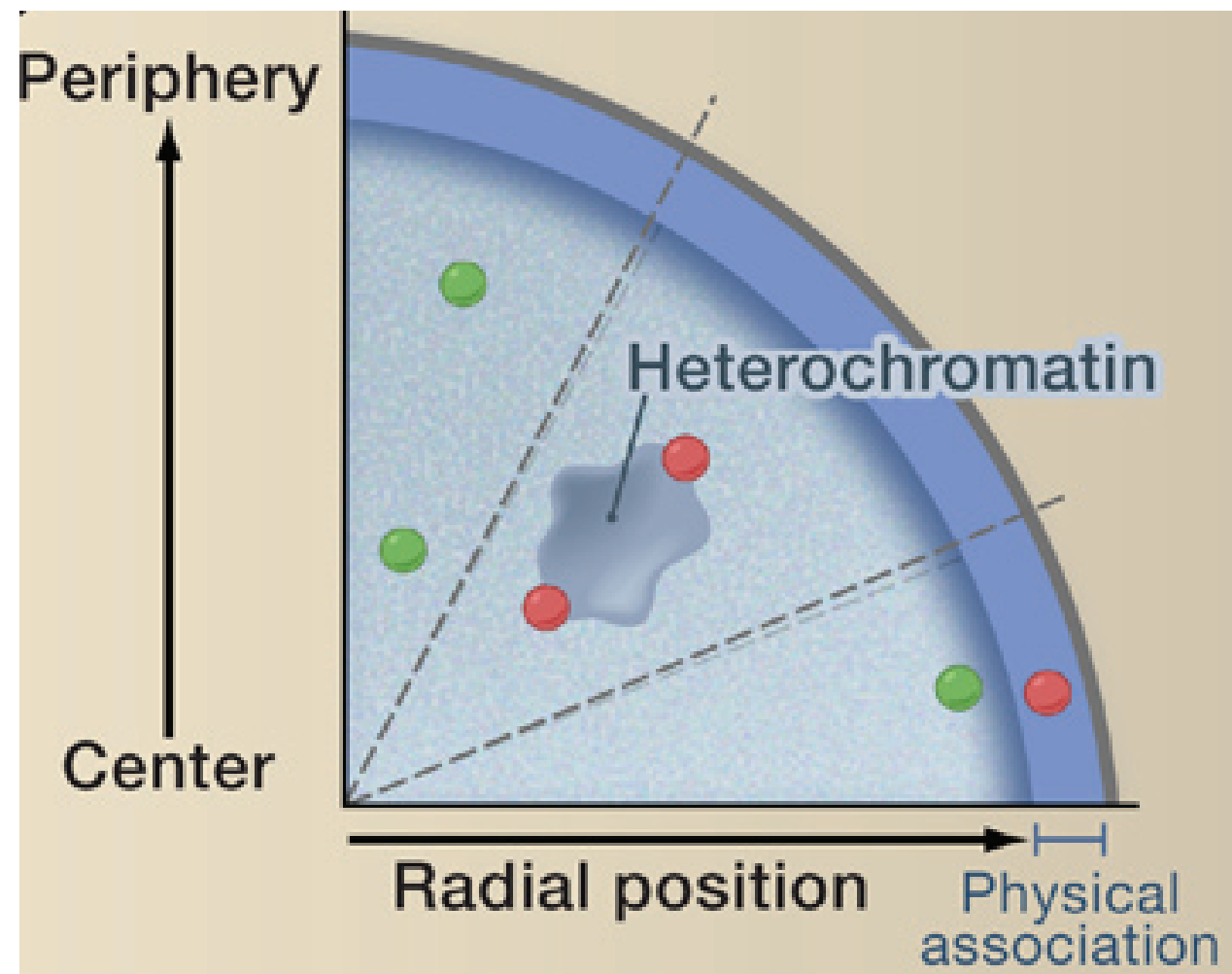
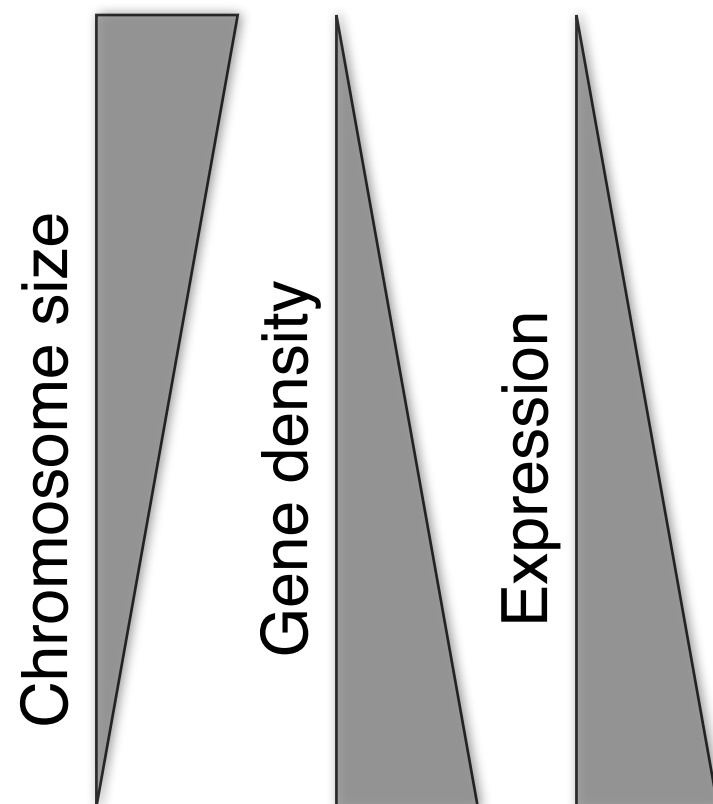
chromatin that is located away from the nuclear lamina, is generally less densely packed, and contains actively transcribed genes

Heterochromatin:

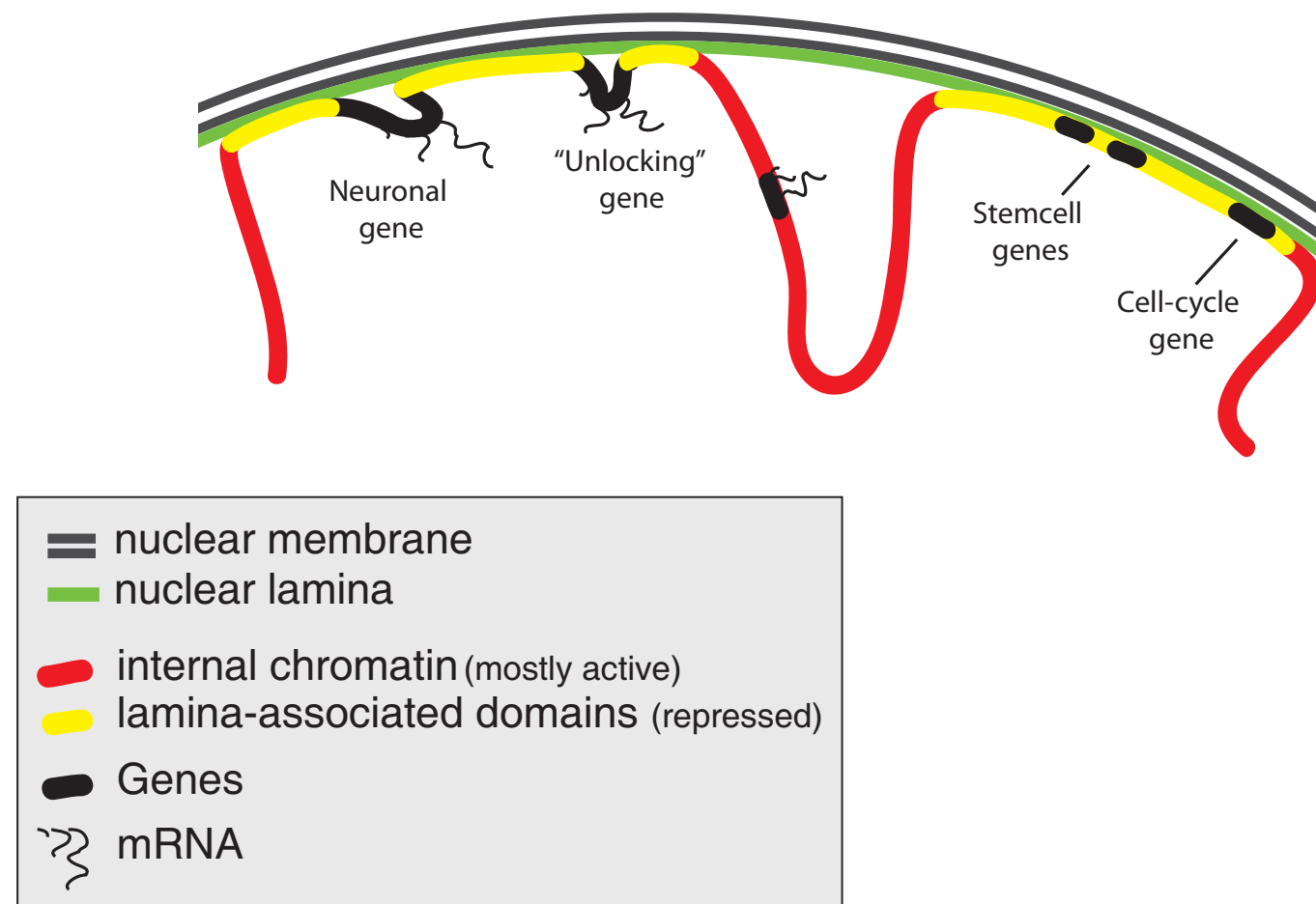
chromatin that is near the nuclear lamina, tightly condensed, and transcriptionally silent

Complex genome organization

Takizawa, T., Meaburn, K. J. & Misteli, T. The meaning of gene positioning. Cell 135, 9–13 (2008).



Lamina-genome interactions

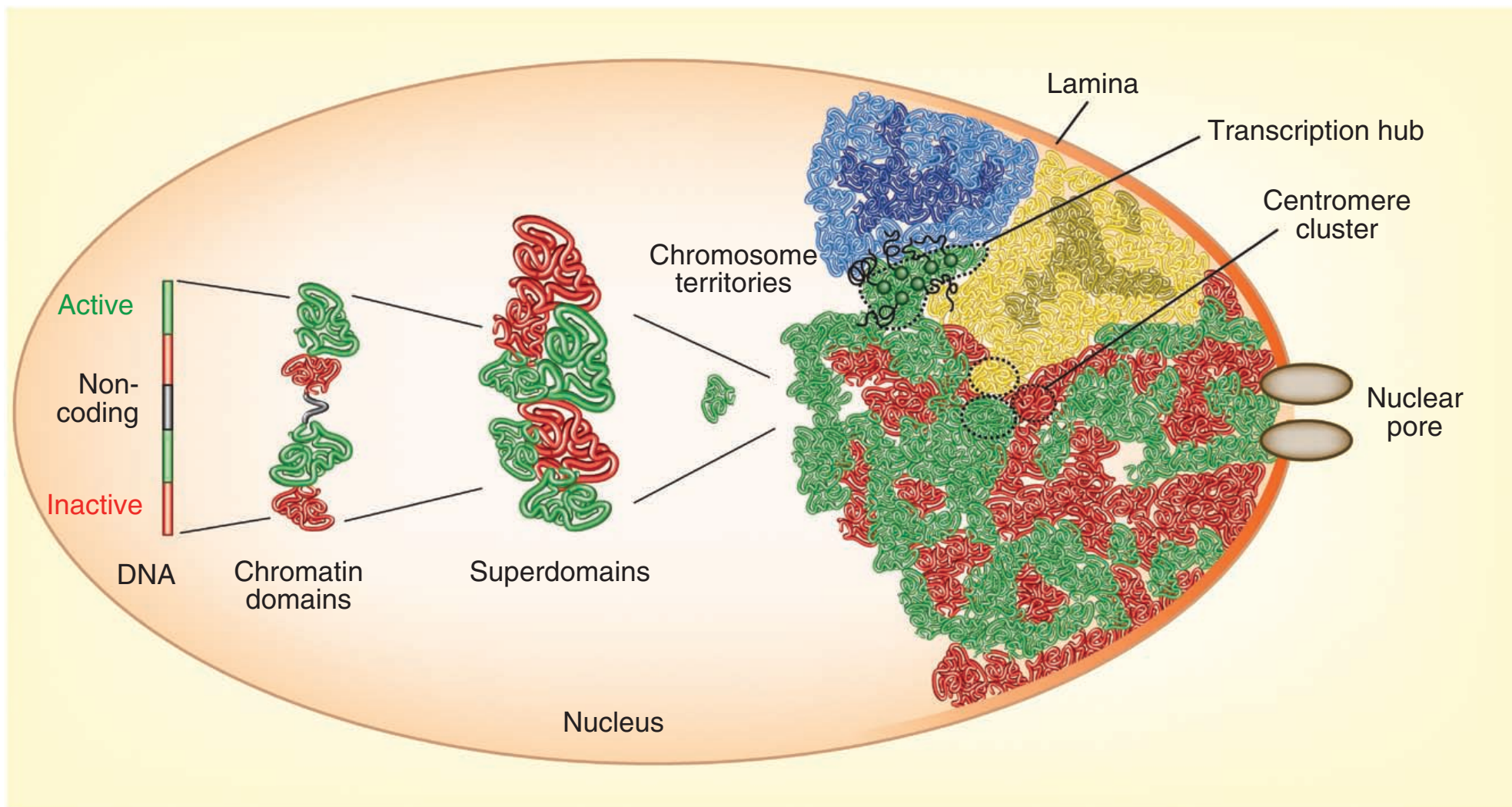


Most genes in Lamina Associated Domains are transcriptionally silent, suggesting that **lamina-genome interactions** are widely involved in the control of **gene expression**

Adapted from Molecular Cell 38, 603-613, 2010

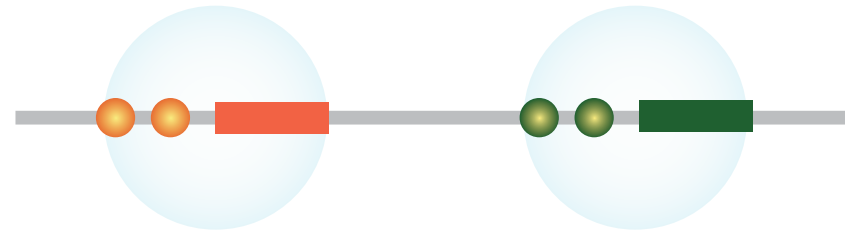
Complex genome organization

Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat Struct Mol Biol* 20, 290–299 (2013).

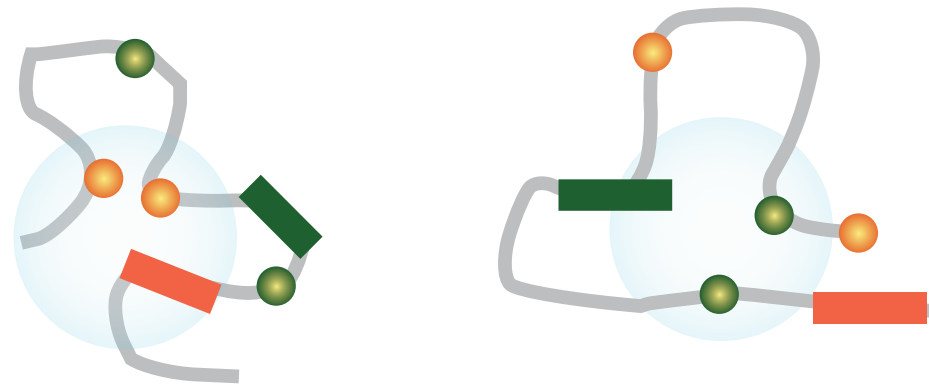


Chromatin loops

Simple



Complex



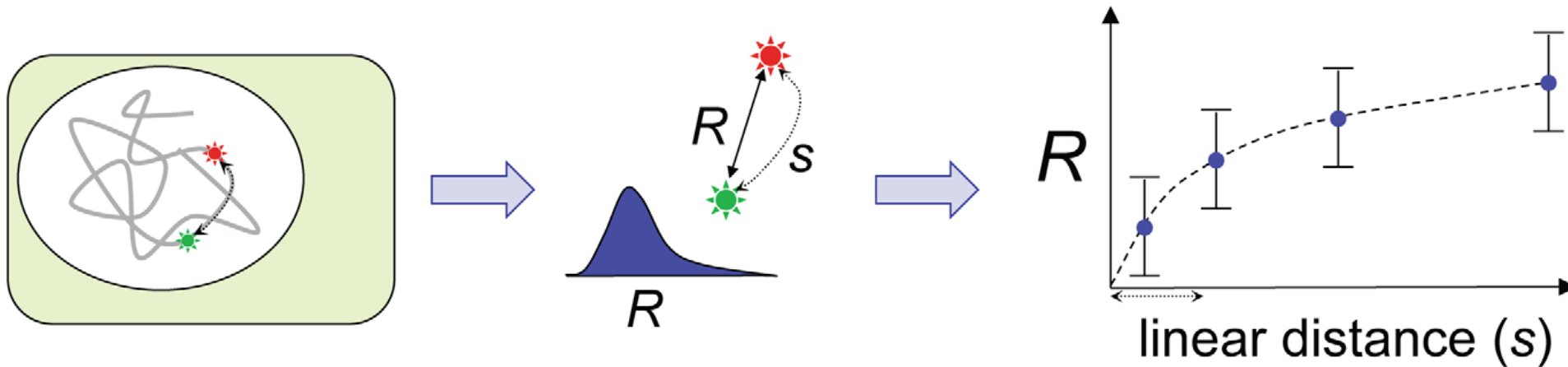
Loops bring distal genomic regions in close proximity to one another

This in turn can have profound effects on gene transcription

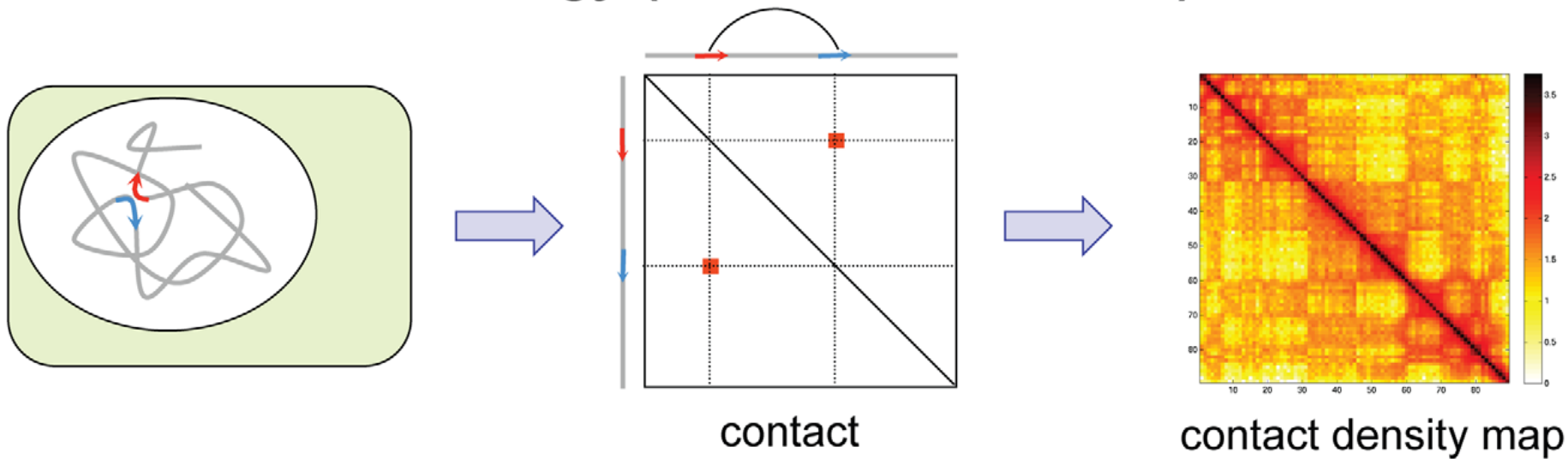
Enhancers can be thousands of kilobases away from their target genes in any direction (or even on a separate chromosome)

Main approaches

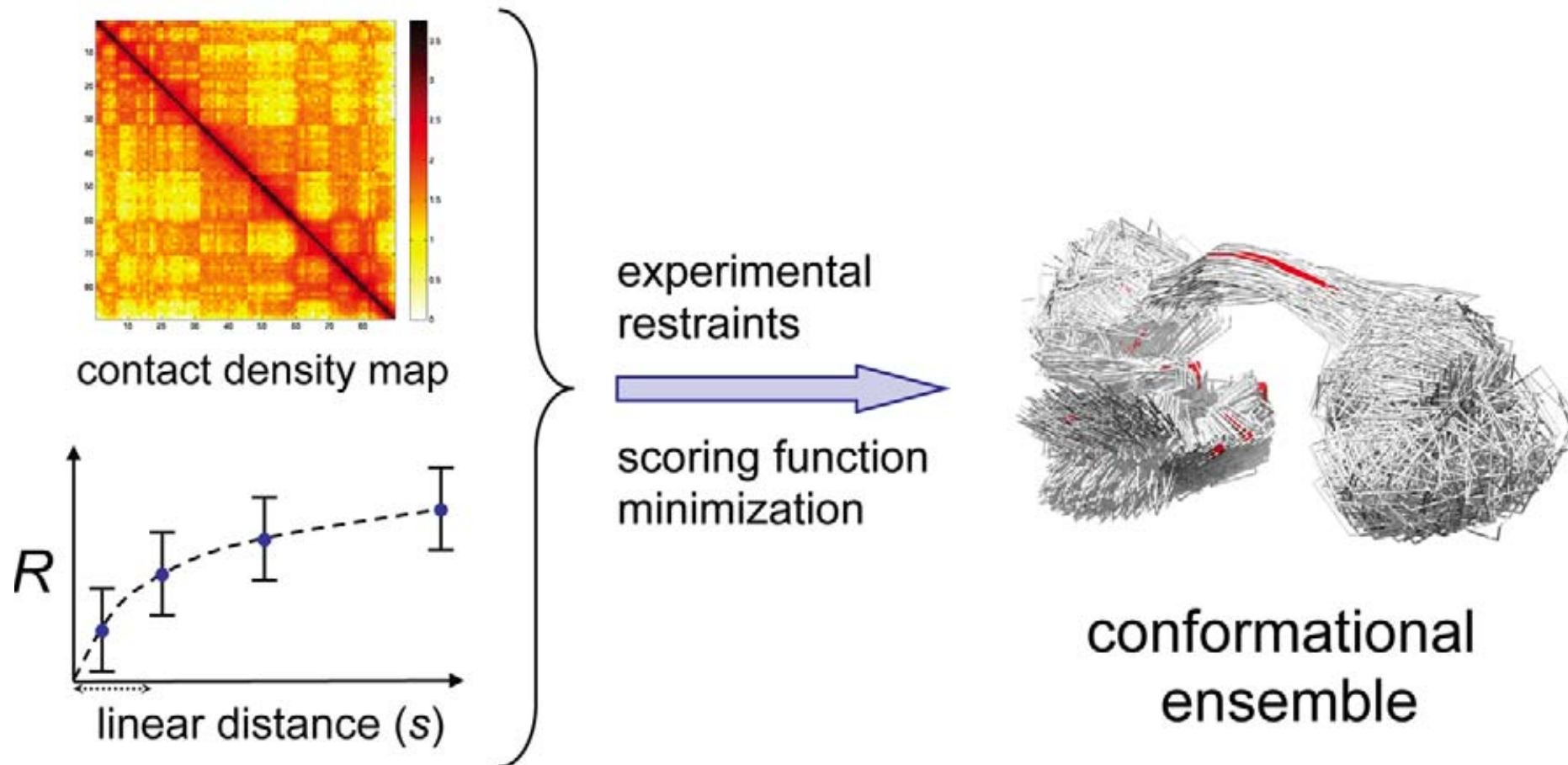
Light microscopy (FISH)



Cell/molecular biology (3C-based methods)



Restrain based modeling (IMP)

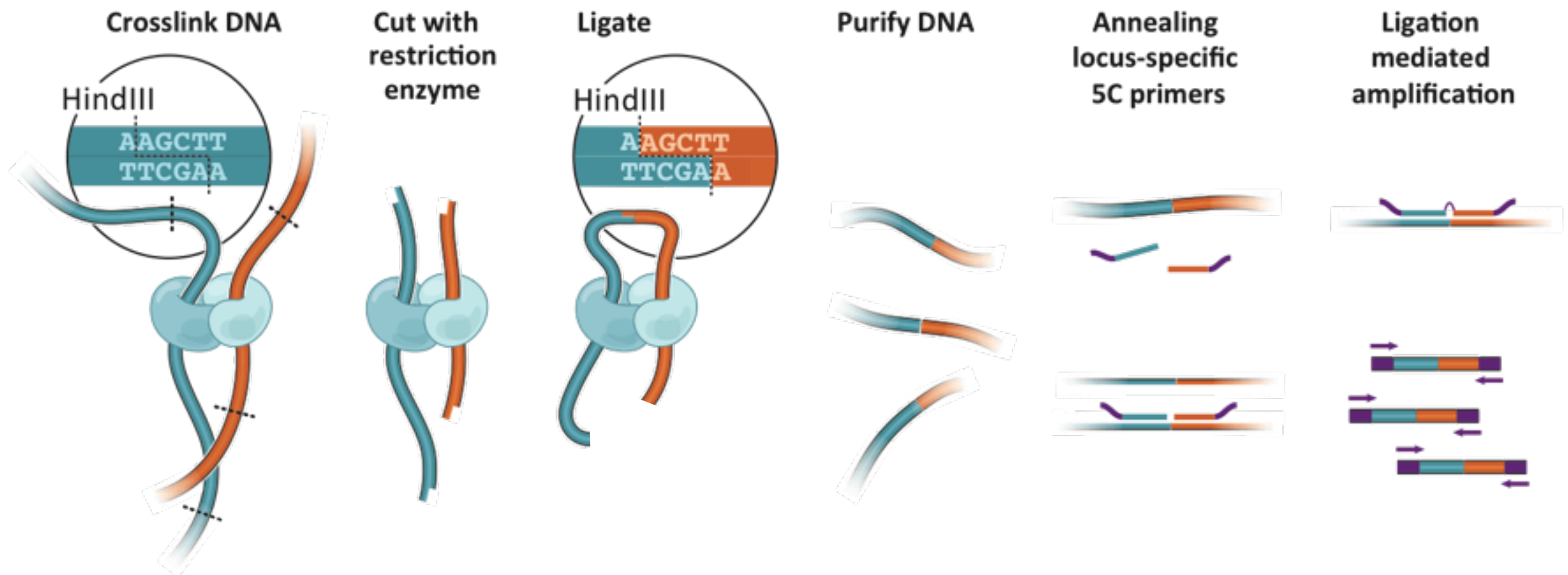


5C technology

<http://my5C.umassmed.edu>



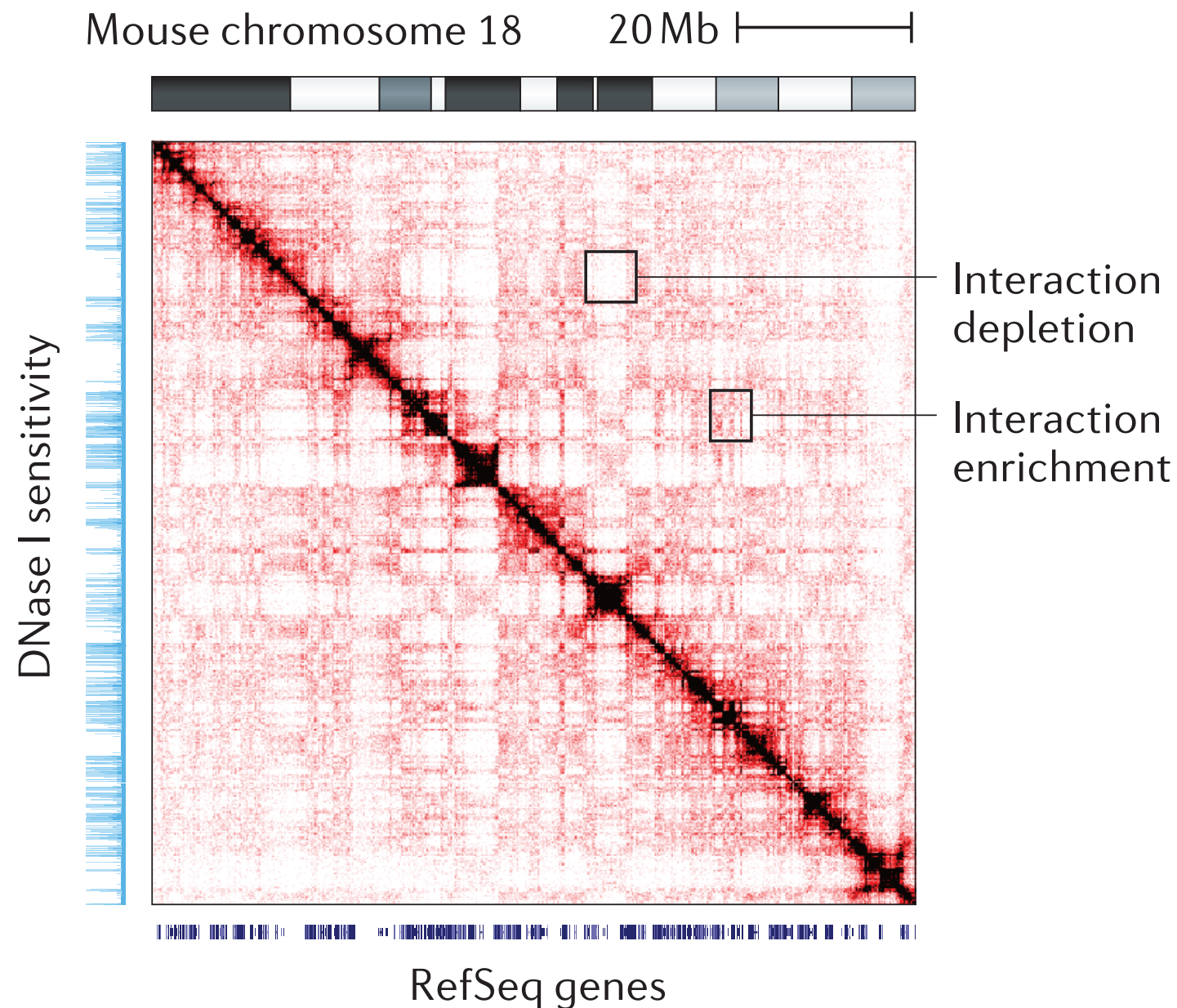
Job Dekker



Dostie et al. Genome Res (2006) vol. 16 (10) pp. 1299-309

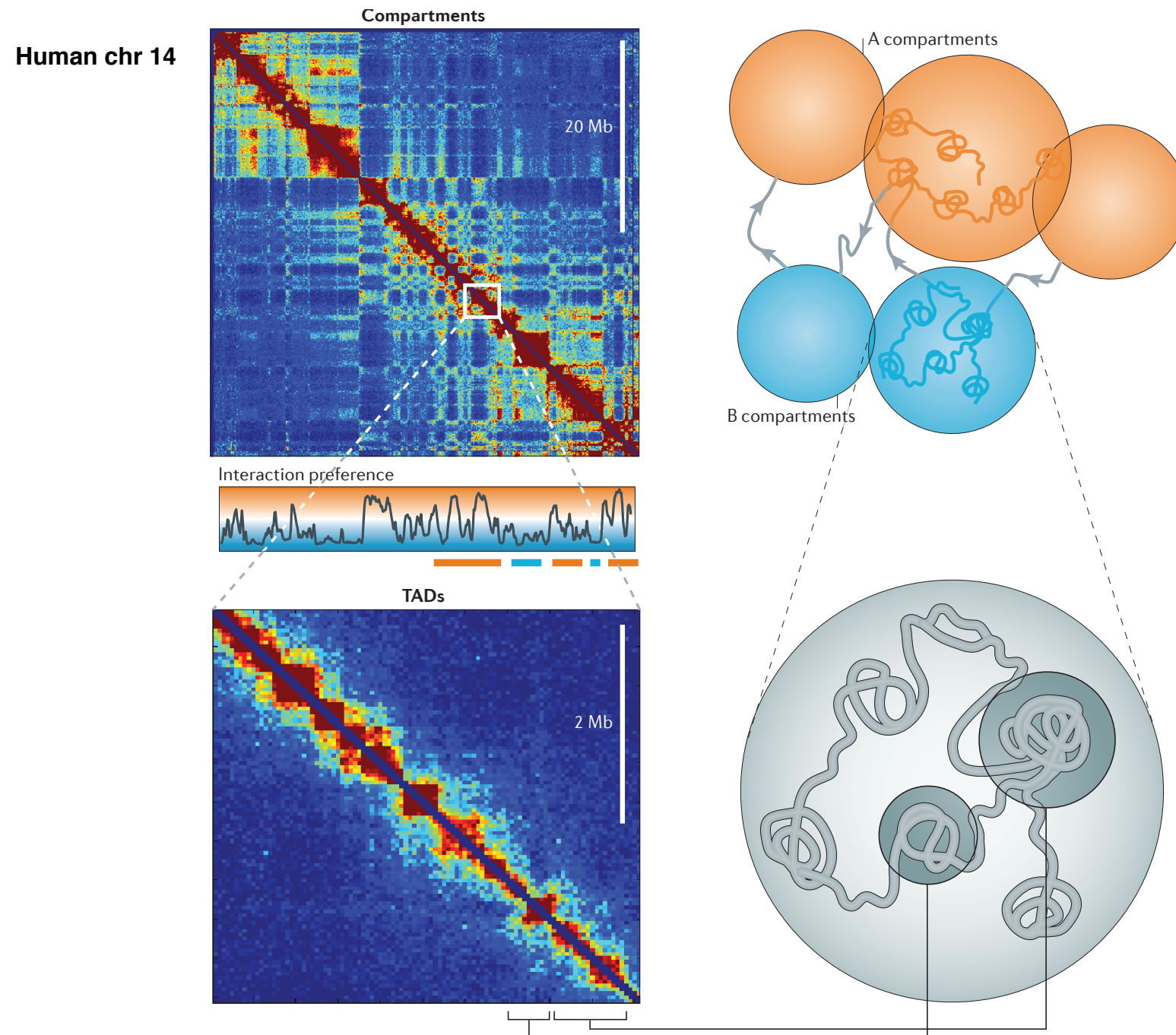
Hi-C data and genomic tracks data

Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14, 390–403 (2013).

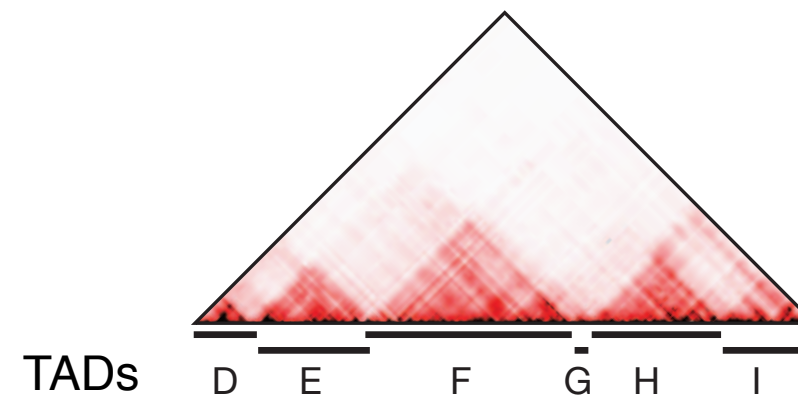
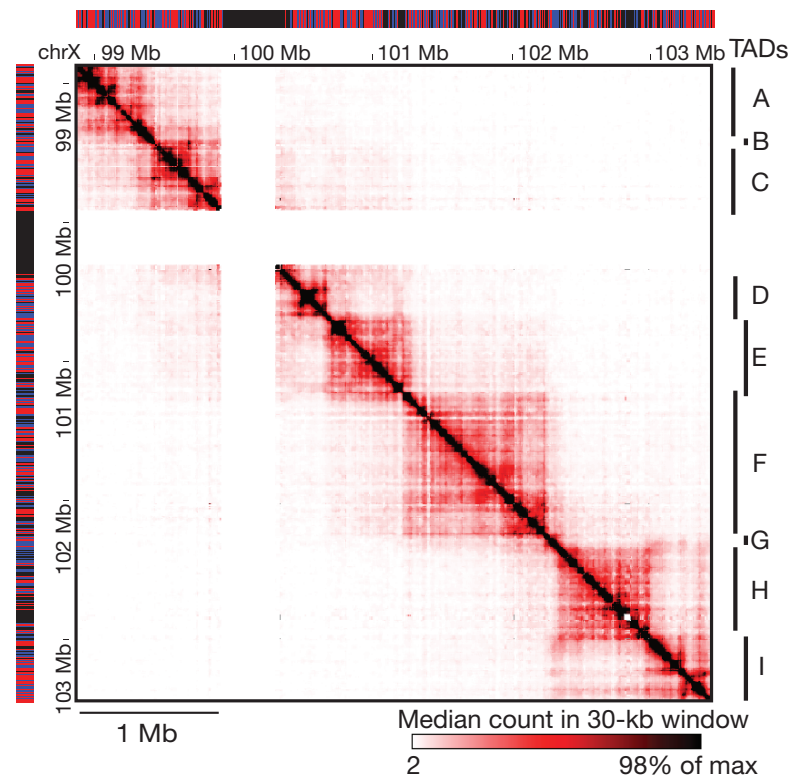


Complex genome organization

Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14, 390–403 (2013).



Topologically Associating Domains (TADs)



Topologically associating domains (TADs) can be made of up to hundreds of kb in size

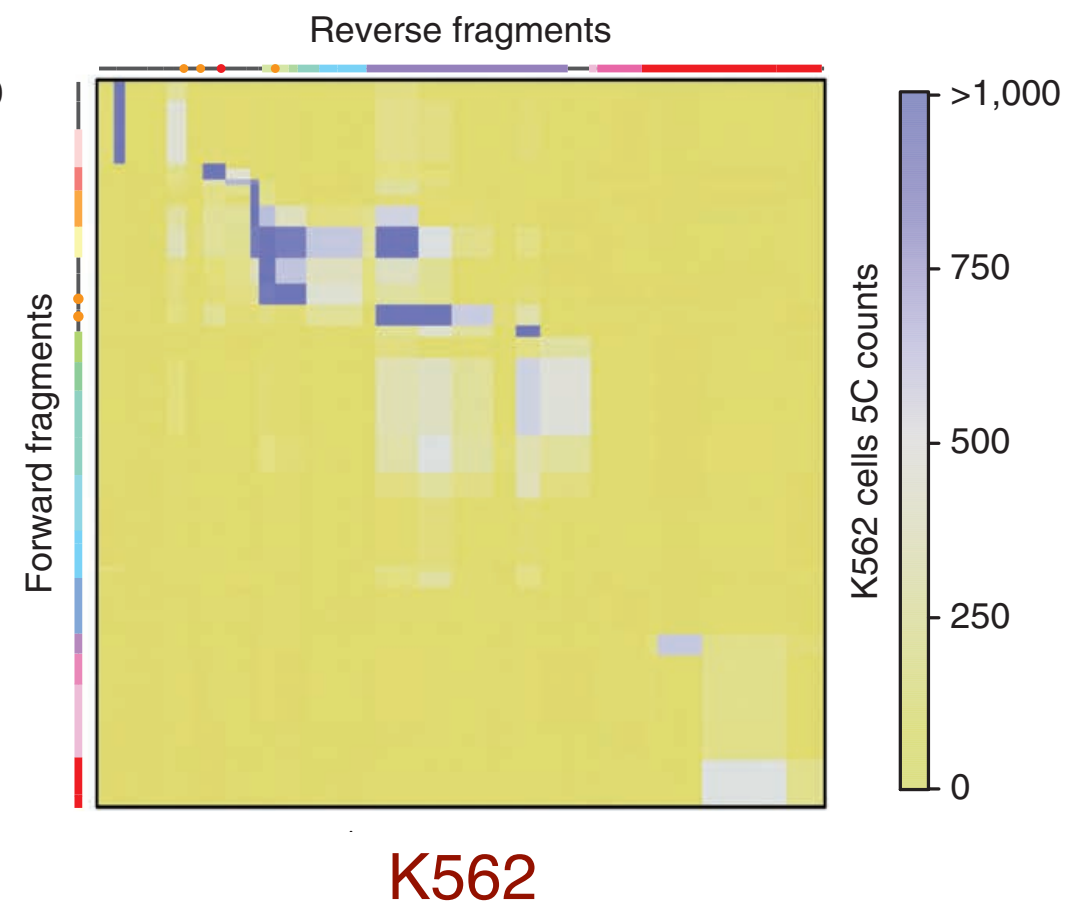
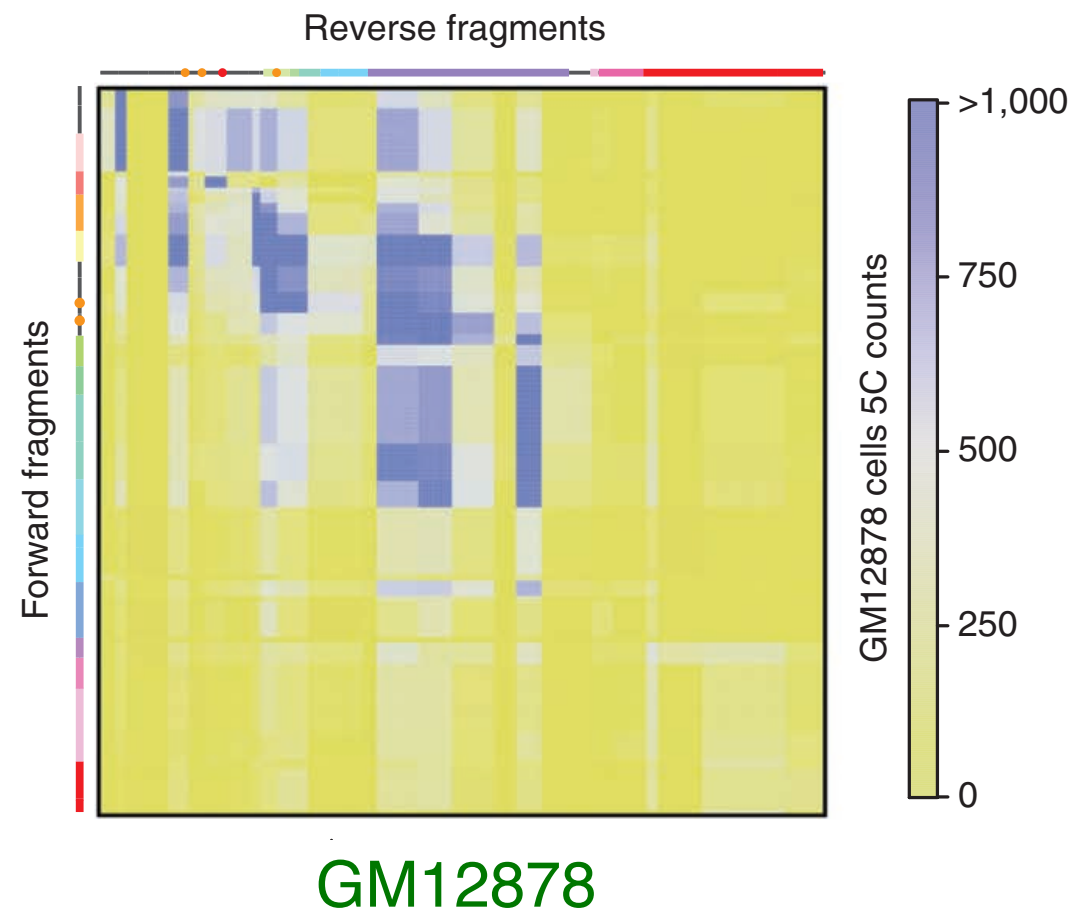
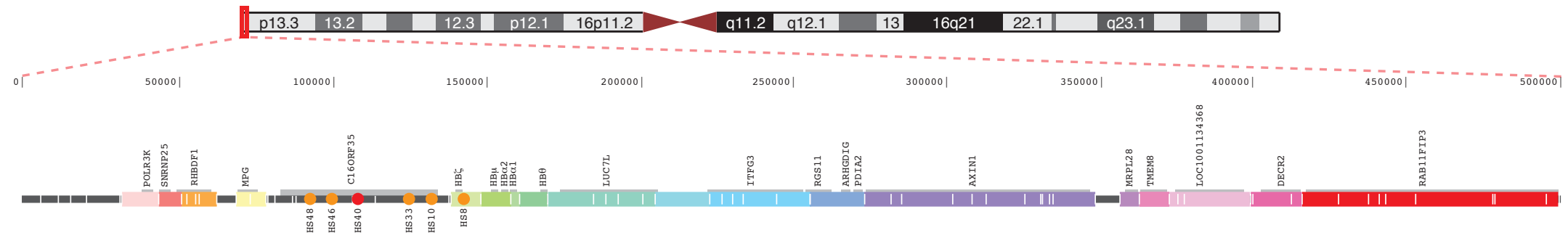
Loci located within TADs tend to interact more frequently with each other than with loci located outside their domain

The human and mouse genomes are each composed of over 2,000 TADs, covering over 90% of the genome

Human α -globin domain

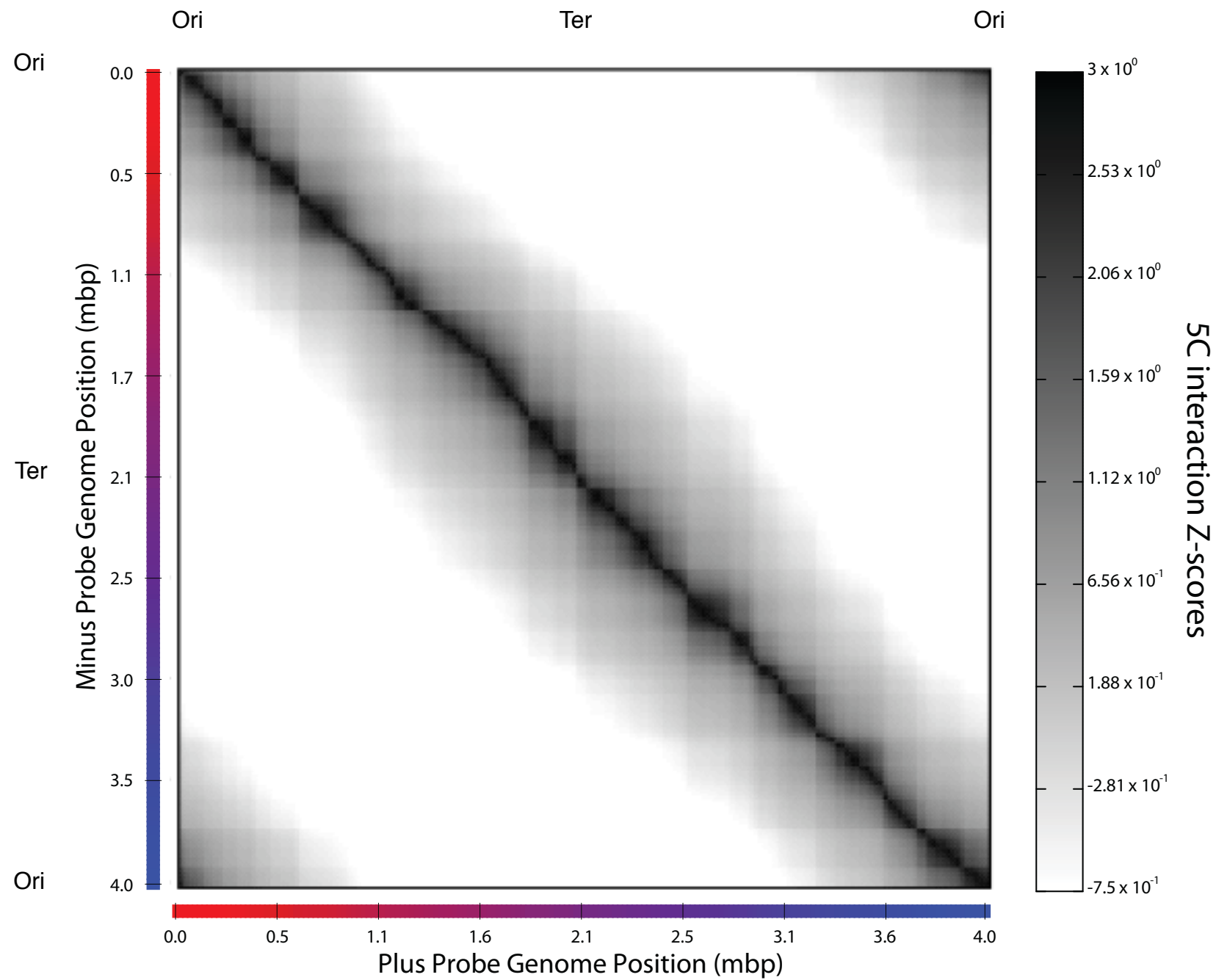
ENm008 genomic structure and environment

ENCODE Consortium. *Nature* (2007) vol. 447 (7146) pp. 799-816



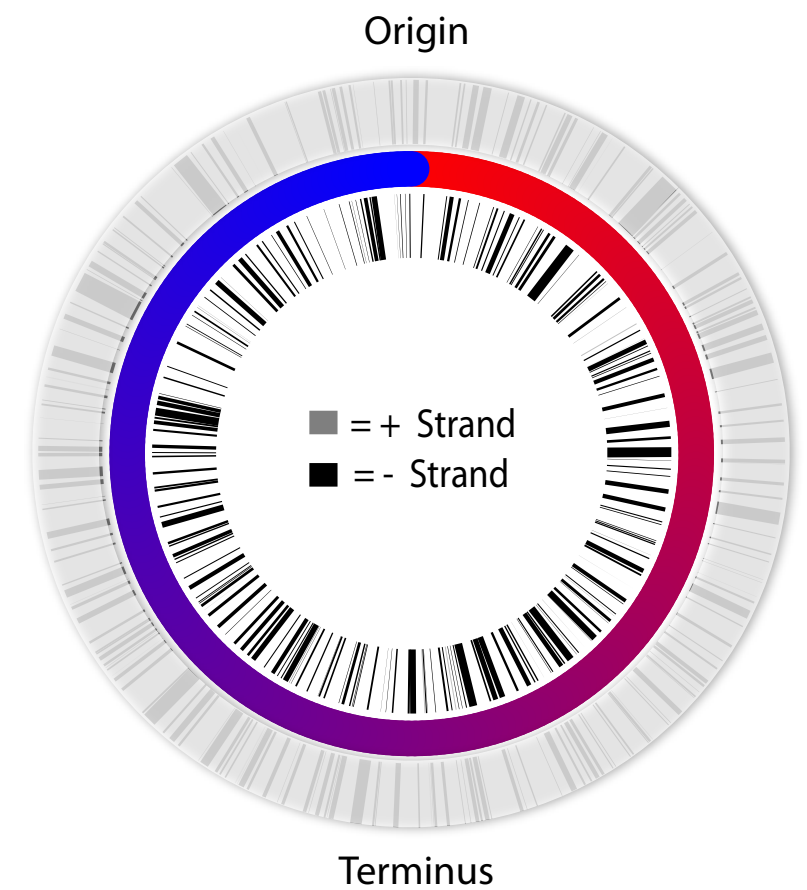
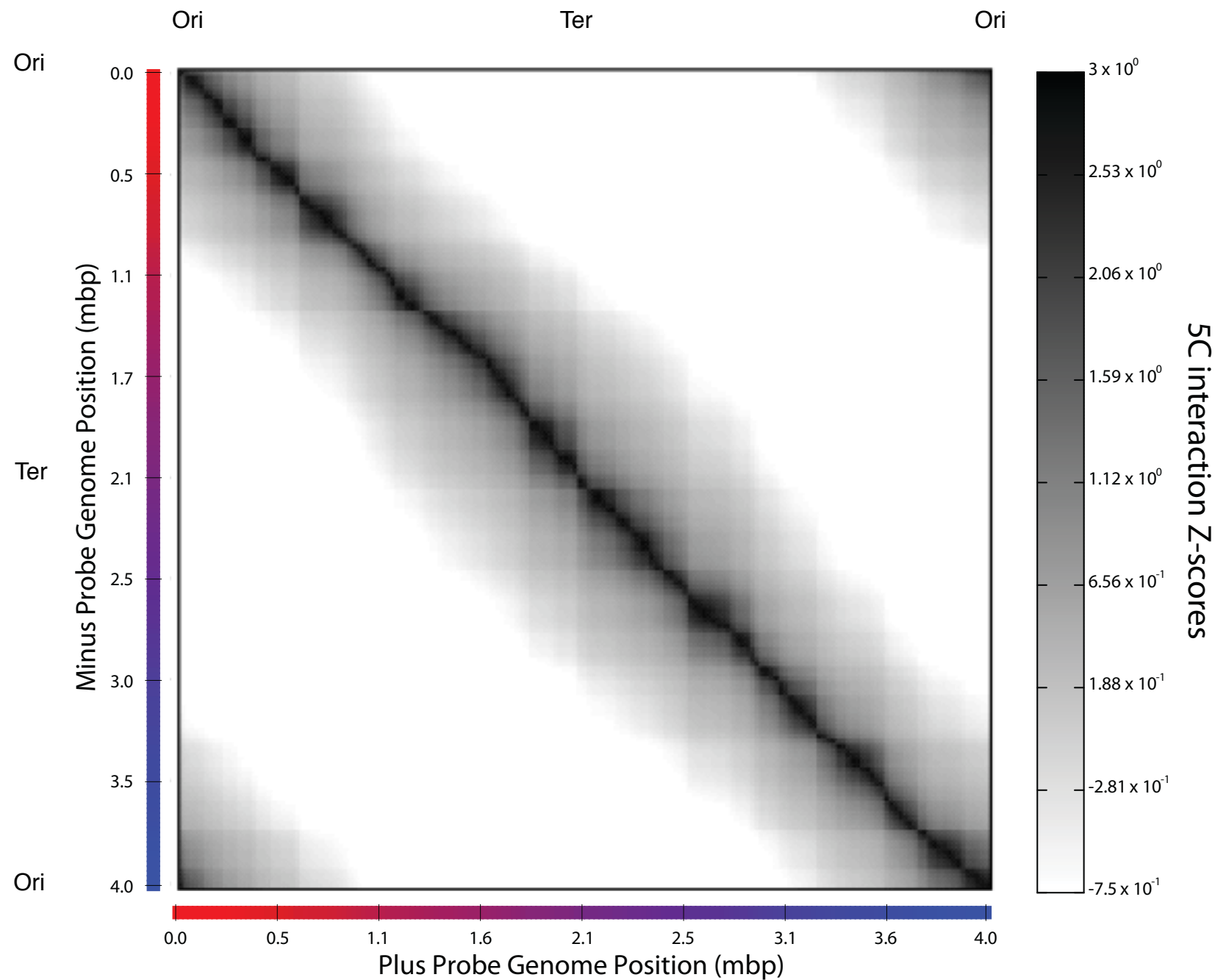
The genome of *Caulobacter Crescentus*

Toy interaction matrix



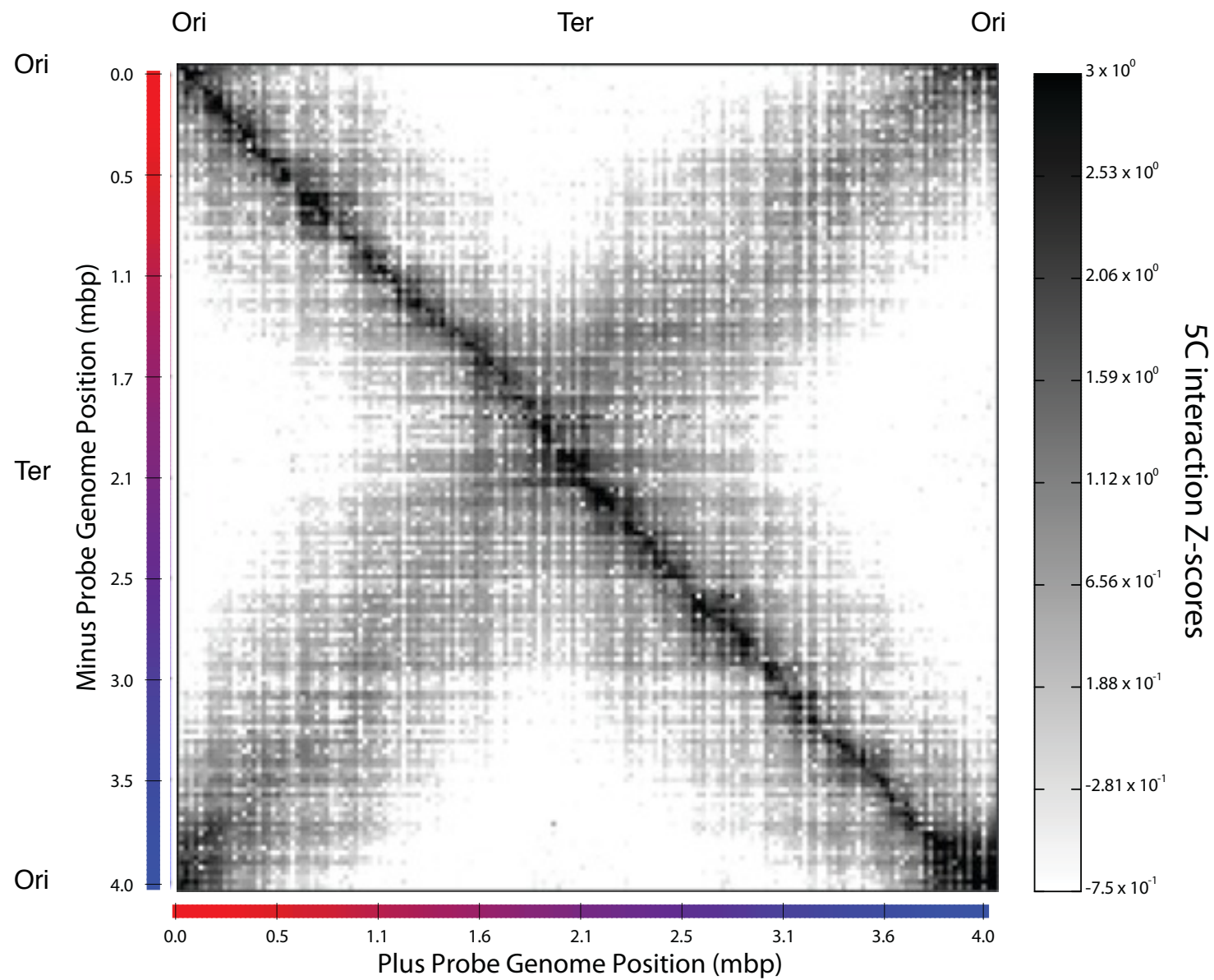
The genome of *Caulobacter Crescentus*

Toy interaction matrix



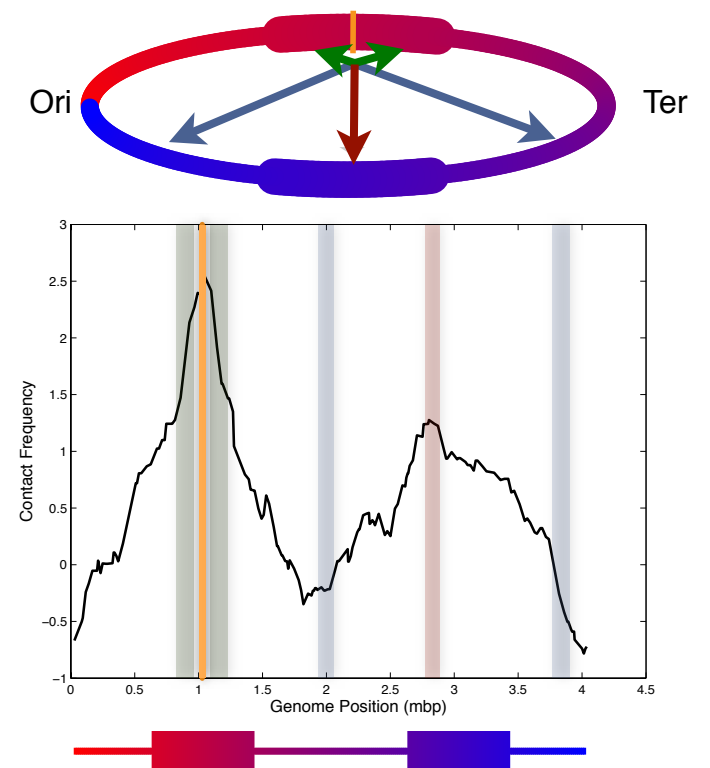
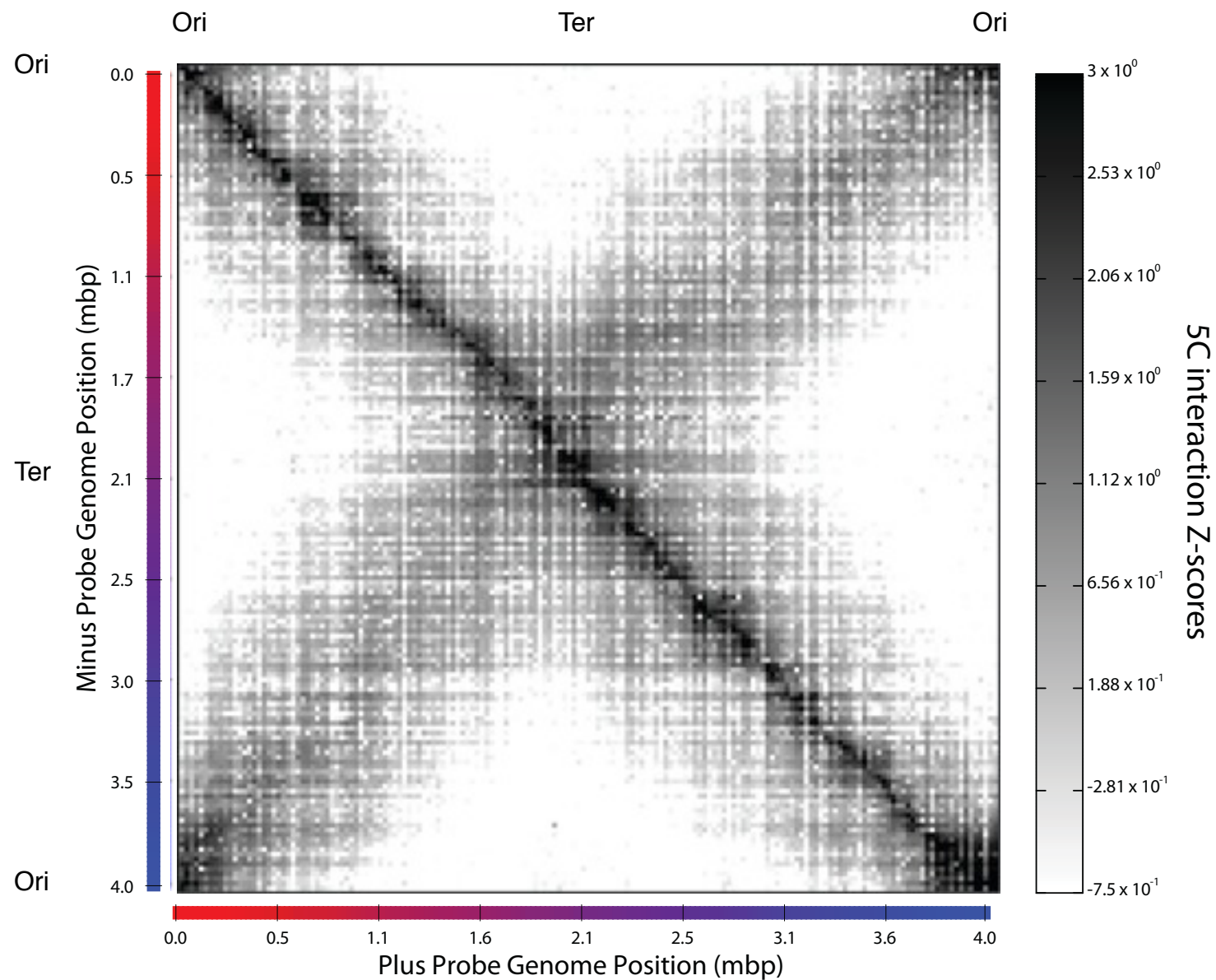
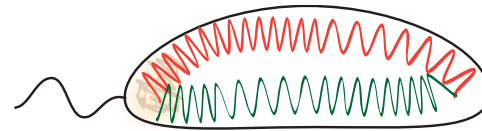
The genome of *Caulobacter Crescentus*

Real interaction matrix



The genome of *Caulobacter Crescentus*

Real interaction matrix



Take home message

Chromatin = DNA + (histone) proteins

The genome is well organized and hierarchically packaged

Histone modifications affect chromatin structure and activity

3C-like data measure the frequency of interaction between distant loci

How DNA is packaged



The Integrative Modeling Platform

<http://integrativemodeling.org>



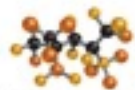
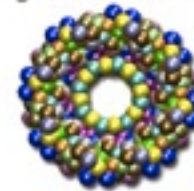
[home](#) [about](#) [news](#) [download](#) [doc](#) [wiki](#) [contact](#)

IMP, the Integrative Modeling Platform

IMP's broad goal is to contribute to a comprehensive structural characterization of biomolecules ranging in size and complexity from small peptides to large macromolecular assemblies, by integrating data from diverse biochemical and biophysical experiments. IMP provides an open source C++ and Python toolbox for solving complex modeling problems, and a number of applications for tackling some common problems in a user-friendly way. IMP can also be used from the [Chimera](#) molecular modeling system, or via one of several web applications.

IMP is open source software, mostly available under the terms of the GNU Lesser General Public License (LGPL). (Some IMP modules are available under the [GNU GPL](#) instead.)

Get started with IMP by [downloading it](#) and [checking out the documentation](#).



NCDIR

The IMP software is used as part of the National Center for Dynamic Interactome Research (NCDIR).

If you use IMP, please cite D. Russel, K. Lasker, B. Webb, D. Schneidman, J. Velázquez-Muriel, A. Sali, "Putting the pieces together: integrative structure determination of macromolecular assemblies", *PLoS Biology*, 2012. The main page of each IMP module in the documentation also lists publications relevant to that module.

Installing IMP

Install the required libraries:

```
sudo apt-get install cmake
sudo apt-get install libboost1.49-all-dev
sudo apt-get install libhdf5-dev
sudo apt-get install swig
sudo apt-get install libcgall-dev
sudo apt-get install python-dev
```

Download the IMP tarball file from <http://salilab.org/imp/> and uncompress it:

```
wget http://salilab.org/imp/get.php?pkg=2.0.1/download/imp-2.0.1.tar.gz -O imp-2.0.1.tar.gz
tar xzvf imp-2.0.1.tar.gz
```

Move to the IMP directory and compile the code

Compiling IMP

```
cd imp-2.0.1
cmake . -DCMAKE_BUILD_TYPE=Release -DIMP_MAX_CHECKS=NONE -DIMP_MAX_LOG=SILENT
make -j4
```

Once the compilation has finished, open the file `setup_environment.sh` in your IMP directory and copy the first lines into your `~/.bashrc` file (if this file is not present in your home directory, create it). These lines should look like:

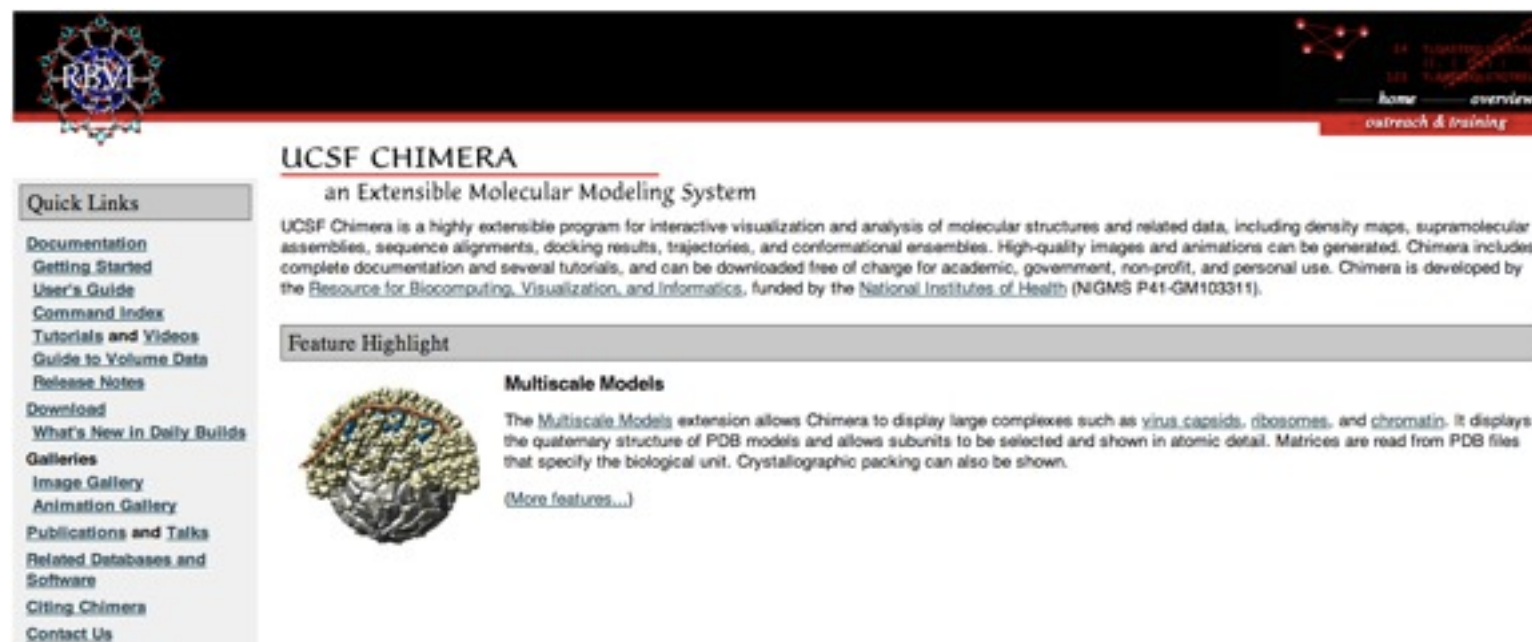
```
LD_LIBRARY_PATH="/SOMETHING/imp-2.0.1/lib:/SOMETHING/imp-2.0.1/lib:/SOMETHING/imp-2.0.1/src/
dependency/RMF/:$LD_LIBRARY_PATH"
export LD_LIBRARY_PATH
```

```
PYTHONPATH="/SOMETHING/imp-2.0.1/lib:/SOMETHING/imp-2.0.1/lib:/SOMETHING/imp-2.0.1/src/dependency/
RMF/:$PYTHONPATH"
export PYTHONPATH
```

**>> Do not copy the lines above, copy them from `setup_environment.sh`,
where **SOMETHING** is replaced by your real path to IMP <<**

Installing Chimera

<http://www.cgl.ucsf.edu/chimera/>



The screenshot shows the UCSF Chimera website homepage. At the top, there is a navigation bar with a logo on the left and links for 'home', 'overview', and 'outreach & training' on the right. Below the navigation bar, the main heading reads 'UCSF CHIMERA' followed by the subtitle 'an Extensible Molecular Modeling System'. A paragraph of text describes Chimera as a highly extensible program for interactive visualization and analysis of molecular structures and related data. To the left of the main content is a 'Quick Links' sidebar with categories like 'Documentation', 'Download', 'Galleries', and 'Publications and Talks'. Below the main heading, a 'Feature Highlight' section titled 'Multiscale Models' includes a small image of a molecular complex and a description of the Multiscale Models extension, which allows displaying large complexes like virus capsids and ribosomes. A link '(More features...)' is provided at the end of the description.

UCSF CHIMERA
an Extensible Molecular Modeling System

UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. High-quality images and animations can be generated. Chimera includes complete documentation and several tutorials, and can be downloaded free of charge for academic, government, non-profit, and personal use. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics, funded by the National Institutes of Health (NIGMS P41-GM103311).

Feature Highlight

Multiscale Models

The [Multiscale Models](#) extension allows Chimera to display large complexes such as virus capsids, ribosomes, and chromatin. It displays the quaternary structure of PDB models and allows subunits to be selected and shown in atomic detail. Matrices are read from PDB files that specify the biological unit. Crystallographic packing can also be shown.

[\(More features...\)](#)

Chimera commands

Align

match.sh #1 #0

Select

select #model:particles

Measure

distance #0:1-2

angle #0:1-2

Display

vdwdefine #radius

shape tube #0 radius 1 bandLength 3 segmentSubdivisions 10

shape tube #0 rad 1 band 3 seg 10

Surface

molmap #all 80

color

transparency